

REMARKS

Reconsideration of the rejections set forth in the Office action dated 3/26/04 is respectfully requested under the provisions of 37 CFR §1.111(b).

Claims 1-20 are pending;

Claims 1, 7-13 and 18-20 stand rejected;

Claims 2-6 and 14-17 are objected to;

Claims 12, 13 and 18 were amended to correctly identify that the secondary set of encoding data is for use by the rewrite module.

I. Rejections under 35 USC §103(a)

Claims 1, 8, 10, 11-13 and 19 stand rejected under 35 USC §103(a) as being unpatentable over the combination of Lopresti et al. (5,748,807) and Nagaishi (6,219,449). This rejection is respectfully traversed as a prima facie case of obviousness has not been presented.

A prima facie case of obviousness is established by one or more references that were available to the inventor and that teach a suggestion to combine or modify the reference, the combination or modification of which would appear to be sufficient to have made the claimed invention obvious to one of the ordinary skill in the art.

A. Overview

One aspect of the claimed invention is described starting at page 41, line 13 and describes a “rewrite model” of an embodiment that uses aspects of the inventive techniques. The preceding portion of the instant application has been directed to inventive techniques for correcting optical character recognition (OCR) errors where the OCR recognition program incorporates an assist channel and a primary channel; where

the assist channel provides the OCR recognition program with additional information that can be used by the OCR recognition program to correct errors or omissions of the character images (page 8, lines 21). Three embodiments for encoding and decoding the assist channel (the separation coding, block coding and convolution coding) have been disclosed prior to page 41. These techniques were described as being incorporated within the OCR decoder module 216 itself (see fig. 2, and page 7, line 18 – page 8, line 2). That is, these descriptions were of embodiments where the OCR function and disclosed use of the assist channel were combined.

The “rewrite model” is directed to embodiments where the OCR function in the decoding module is independent of the assist channel correction function. Thus, an off-the-shelf OCR program (decoding module) can be used to recognize characters in the primary channel and the output of the decoding module output (the candidate primary channel) can be adjusted using the assist channel independently from the decoding module (page 41, lines 22-28). One aspect of the rewrite model is an event library that identifies potential OCR error modes and provides correction rules and weightings for those error modes (page 42, lines 7-20). The rules in the event library coupled with the information in the assist channel allows this aspect of the invention to detect and repair decoding errors in the candidate set of symbol data generated by the OCR module.

This aspect of the invention is claimed by **claims 1, 10, and 11**. Claim 1 is:

1. A method for decoding image data for a hardcopy document, comprising:

recording a scanned representation of the hardcopy document that includes a primary set of symbol data and a secondary set of encoding data; the primary set of symbol data providing a first channel of human readable information rendered on the hardcopy document; the secondary set of encoding data providing a second channel of machine readable information rendered on the hardcopy document;

receiving a decoded form of the scanned representation of the hardcopy document from a decoding module to define a candidate set of symbol data; and

rewriting, independent of the decoding module, the candidate set of symbol data using an event library and the secondary set of encoding data; the event library identifying likely failures encountered when the scanned representation of the hardcopy document is decoded.

B. Cited Art

Lopresti: The Lopresti reference discloses techniques for adding additional information to a page that will be scanned. When the page is scanned the OCR conversion can use the additional information (for example, to detect and correct errors in the conversion). Lopresti does not teach rewriting the candidate set of symbol data using an event library and the secondary set of encoding data. Nor does Lopresti teach the event library identifying likely failures encountered when the scanned representation of the hardcopy document is decoded. Furthermore, nothing in Lopresti teaches a suggestion to make such a combination.

Nagaishi: The Nagaishi reference uses the term dictionary in the background of the invention section of his patent to 1) specify which characters are likely to have recognition errors, and 2) to specify a dictionary related to grammar against which a combination of possible characters can be checked for correctness (C1, lns. 30--44). In addition, Nagaishi discloses techniques for matching hand-written characters with a library of character shapes. Nagaishi also discloses techniques for recognizing semantic units such as words, lines, sentences, or paragraphs (C:3, lines7-11). This is done by searching for combinations of characters in the semantic unit (for example, a word) until a valid combination of characters is found (in the word case, until a combination of characters matches a word search from a dictionary). Nagaishi's improvement is to

adjust the weighting of the possible characters to reduce the number of combinations that must be searched to find a match. The weighting is determined by statistically analyzing hand written samples to determine the variation in the character depending on where the character is placed in the semantic unit (C:3, lns. 14-30). When a string of characters are first analyzed in a word, the characters in the character groups that have the highest probability of variation are selected before the characters with the lower variability and the results of each selected combination is checked against a dictionary. The combination is used if the combination is a valid word in the dictionary (C:5, ln. 63 – C:6, ln. 36). By first selecting characters from groups with the higher probability of variation, Nagaishi tends to increase the probability of finding a combination of characters that have a grammatical meaning.

Applicant respectfully traverses the office action assertion that Nagaishi's dictionary is an event library. The office action asserts that the characters in Nagaishi's dictionary are events, and the dictionary is a library.

As discussed above, the event library along with the secondary set of encoding data is used to detect and correct an error made by the decoding module. The event library consists of rules that specify transformations if the secondary set of encoding data (the assist channel data) indicate that an error exists in the candidate set of symbol data. The rules in the event library are triggered on the detection of an event (that being the detection of an inconsistency between the information in the assist channel and a symbol in the candidate set of symbol data. Once an event is triggered, a rule from the event library is selected responsive to the inconsistent character and the corresponding information in the assist channel.

Nagaishi's character dictionary is simply a grouping of possible characters that may match a handwritten pattern and a probability adjustment depending on where the handwritten pattern is located with respect to other handwritten patterns. Nagaishi's other dictionary is a dictionary of grammatical meanings to which possible combinations of characters from the character dictionary may match. Neither of Nagaishi's dictionaries include rules or transformations. Thus, Nagaishi does not teach the event library nor does it teach a suggestion for a modification to add the event library.

Furthermore, nothing in Lopresti and Nagaishi would suggest the combination to one skilled in the art. Thus, claims 1, 10 and 11 are patentable.

With regard to **claims 8 and 19**: These claims depend on and further limit their respective parent claims that are patentable. Thus, claims 8 and 19 are patentable. In addition, applicant agrees that the separation coding, block coding, and convolution coding are well known error correction mechanisms. However, applicant fails to understand the Office Action's assertion that Lopresti's secondary encoding data can be a special font, bar code or other symbology. The information in the assist channel can be recorded in any manner on the document in any fashion. However, the information in the assist channel can be developed separation coding, block coding, and convolution coding and thus, the information in the assist channel can be used to correct errors in the candidate set of symbol data,

Currently amended **claims 12 and 13** depend on and further limits claim 11 that is patentable. Thus, claim 12 and 13 are patentable. In addition, with respect to claim 13, applicant traverses the assertion that a barcode is an inherently compressed form of data.

Claims 7, 9, 18 and 20 stand rejected under 35 USC §103(a) as being unpatentable over the combination of Lopresti et al. (5,748,807) and Nagaishi (6,219,449) and Kopec (5,594,809). This rejection is respectfully traversed as a prima facie case of obviousness has not been presented.

Claims 1 and 11 are patentable.

Claims 7, 9, 18 and 20 depend from and further limit their respective independent claim. Thus claims 7, 9, 18 and 20 are also patentable.

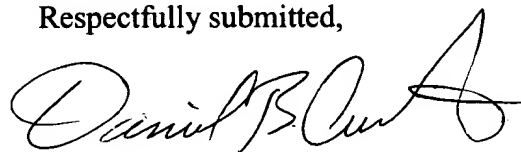
No additional fee is believed to be required for this amendment as the three month shortened statutory period ended on a Saturday and this office action reply was sent to the USPTO on the following Monday. However, the undersigned Xerox Corporation authorized attorney hereby authorizes the charging of any necessary fees, other than the issue fee, to Xerox Corporation Deposit Account No. 24-0025. This also constitutes a

PATENT

request for any needed extension of time and authorization to charge all fees therefor to Xerox Corporation Deposit Account No. 24-0025.

Should any additional issues remain, or if I can be of any additional assistance, please do not hesitate to contact me at (650) 812-4259.

Respectfully submitted,

A handwritten signature in black ink, appearing to read "Daniel B. Curtis", with a stylized flourish at the end.

Daniel B. Curtis
Attorney for Applicants
Reg. No. 39,159
(650) 812-4259
dbcurtis@parc.com